DOCUMENT RESUME

ED 251 792 CS 007 827

AUTHOR Perkins, Dorothy
TITLE Assessment of the Use of the Nelson Denny Reading Test.
PUB DATE 84
NOTE 8p.; "Forum for Reading" is the journal of the College Reading Improvement Special Interest Group of the International Reading Association.
PUB TYPE Information Analyses (070)
JOURNAL CIT Forum for Reading; v15 n2 p64-69 Spr-Sum 1984

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Higher Education; *Reading Comprehension; Reading Diagnosis; *Reading Tests; Screening Tests; *Testing Problems; *Test Reviews; *Test Use; *Test Validity
IDENTIFIERS *Nelson Denny Reading Tests

ABSTRACT

The Nelson Denny Reading Test (NDRT) is probably the most widely used test of reading comprehension at the college level in the nation. However, reviews of the test, as well as recent reports of its failure to adequately measure gain or lack of gain of college students enrolled in reading improvement courses, do not support the popularity it has enjoyed. Reviewers of the NDRT have criticized the test for its difficult passages, stringent time limitations, and poor normative data. The conclusions drawn by some reviewers of the NDRT are that the test may have value at the college level when used for screening purposes, but that care should be exercised in interpreting scores for purposes of individual diagnosis. Studies of other reviewers support these conclusions. In fact, some suggest that the use of the NDRT on anyone other than highly selected college students or adults would produce very questionable measurement results and provide a difficult, defeating, and frustrating experience to many students. Thus, a reexamination of the use of the NDRT as the instrument that determines whether or not college freshmen possess adequate reading skills is strongly needed. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED251792

Assessment of the Use of the Nelson Denny Reading Test

Dorothy Perkins

FORUM FOR READING

Volume 15 **Spring/Summer 1984** **Number 2**

cs207827

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Rona F. Flippo

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Assessment of the use of the Nelson Denny Reading Test

Dorothy Perkins
Salisbury State College
Salisbury, Maryland

The Nelson Denny Reading Test is probably the most used test of reading comprehension at the college level in the nation today (Gordon & Flippo, 1983). However, reviews of the test, as well as recent reports of its failure to adequately measure gain or lack of gain of college students enrolled in reading improvement courses, do not support the popularity it has enjoyed (Flippo, 1980).

Reviewers of the NDRT have criticized the test for its difficult passages, stringent time limitations, and poor normative data. According to Raygor (1978, p. 1210), "The test was developed on a very high-level population in which 87 percent of the sample were above the status of college freshmen and 64 percent were upperclass college students, graduates, and faculty," resulting in "...very difficult, bookish passages followed by highly artificial questions of the sort that one would not ask in real life."

Orr (1965, pp. 1077-8), in his review of the NDRT, stated that he "...feels the essence of reading comprehension lies more in the reader's ability to grasp ideas and their implications than in his ability to thread his way through grammatical complexities."

The use of the content described as "very difficult, bookish passages" (Raygor, 1978, p. 1210) which are "restricted ...(to content) chiefly general and somewhat literary in character" (Townsend, 1965, p. 1079) may be supported by reasoning that college students will be reading that type of material in their textbooks, but there is no support for placing severe time restrictions on the reading of passages which should be read with a study-reading approach rather than skim-and-guess.

In addition to content and time restrictions, "...the normative data supplied for grades 13 through 16 do not permit any very descriptive norm-referenced interpretations of test scores. The college level norms cannot be considered as representing any well-defined population." (Forsyth, 1978, p. 1208). The authors of the test admit to problems in obtaining information in establishing college norms:

> In spite of the diligent efforts of the publisher's representatives, cooperation was elicited from a very small portion of the desired sample...The results presented here are, therefore, not presented as a representative sample of the national college population. (Brown, Nelson & Denny, 1976, p. 31)

The authors of the test (Brown et al., 1976, p. 3, 14) suggest that the NDRT is useful for predicting academic success. Reviewers of the NDRT, however, question its predictive validity. "The manual (for forms C and D) indicates that in one sample the NDRT was 'slightly better' than the *Minnesota Scholastic Aptitude Test* in prediction, when the two zero-order correlations involved were .397 and .384, respectively, and both were lower than high school rank at .585." (Raygor, 1978, p. 1209) "...(It) seems reasonable to conclude that the Nelson-Denny adds little to our ability to predict grades beyond what we can do with high school rank alone." (Forsyth, 1978, p. 1208)

BEST COPY 64

Orr (1965, p. 1078), reviewing NDRT forms A and B, found the sparcity of predictive validity one of "...the most disappointing aspects of the present manual," and Cites (1965, p. 108) deplored its unverified statements on predictive validity.

Forms C and D are purported to be "...slightly different subject matter and slightly different levels of difficulty...(to) provide greater latitude for users..." (Brown et al., 1976, p. 3) from forms A and B, but Raygor (1978, p. 1209) suggests that information given on predictive validity seems to be related to previously published forms. Research cited, in fact, was obtained from a review of factors related to success in college published in 1949!

Programs that use NDRT scores for placement in remedial reading programs should be warned against grouping students homogeneously according to their grade equivalent scores. This practice, while convenient, is an imprudent use of grade equivalents. Brown et al. (1976) warns:

> Although easy to understand, grade equivalents must be interpreted with caution, particularly at the upper levels. Among other things they assume a regular pattern of growth throughout the school year, a condition which may not be met. Furthermore, in the area of reading, rather wide deviations should be considered normal. (p. 31)

Morever, 80 percent of the grade equivalents presented on the NDRT (grades 6.0-9.1 and 12.4-15.0) are extrapolated!

The conclusions drawn by reviewers of the NDRT (Cummins, 1981; Forsyth, 1978; Orr, 1965; Raygor, 1978; Townsend, 1965) are that the test may have value at the college level when used for screening purposes, but that care should be exercised in interpreting scores for purposes of individual diagnosis.

Studies carried out by Cummins (1981) and Van Blerkom and Van Blerkom (1981) support the conclusions of the reviewers. Van Blerkom and Van Blerkom questioned students' increased posttest scores in vocabulary and comprehension following instruction in improving reading rate. They found that their students had increased their scores in vocabulary, comprehension, and rate but decided that "...the increases in vocabulary and comprehension were more apparent than real, due to the fact that with increased reading rate the students were able to attempt more items on the second testing," the mean number of items completed having increased significantly ($p < .001$). "However the percentage of attempted items that were correct decreased slightly (not significantly) from the pretest to the posttest."

Cummins (1981, p. 57) compared the percentage of comprehension questions answered correctly of students who were given normal administration of form E of the NDRT and immediately following were given form F with instructions to read the questions and locate the answers, never completely reading the passages. The percent of correct answers increased. This improvement is "...actually an artifact of test format familiarity or a new test marking technique." Cummins also noted that "Low vocabulary and comprehension scores may not always reflect a subject's low achievement in these areas but may be the result of slow cognitive processing, perceptual deficits, or test anxiety."

Raygor and Flippo (1981) used the NDRT and the SDRT to establish criterion related validity for the Minnesota Reading Assessment (MRA) and concluded that the SDRT "is a good group diagnostic test and does

BEST COPY

discriminate well for this population (community colleges, business schools, technical schools and other post-secondary institutions)", and the MRA is a survey test which "identifies students having real problems with weak skills." However, the NDRT "is not an appropriate test to identify students who need help with reading skills in a less than adequately prepared group...(It) is excellent for identifying very skilled readers."

Similar problems were encountered when the NDRT was chosen by Maryland state colleges as the instrument which determines if a college freshman needs remediation in reading and if these students remediated have obtained the necessary comprehension skills for success in college. At Salisbury State College (SSC) it was determined that the use of the NDRT led to errors in the identification of target students, that the scores provided no accurate diagnostic information on which to base remediation, and that students whose reading comprehension was weak could reach the required exit level by applying good test taking techniques. Therefore, additional reading tests have been given to SSC students who have failed to reach the required level (RS 32, 22nd percentile) on NDRT form C.

In the fall of 1980 the Reading Comprehension subtest of the *Descriptive Test of Language Skills* (DTLS) was given to all students at SSC who failed to reach the required level on the NDRT. Table 1 presents the DTLS mean and range of scores received by students at each possible NDRT score. For example, 33 students received a raw score of 32 (22nd percentile) on the NDRT. Those 33 students had a mean raw score of 33.1 (47th percentile) on the DTLS. The range of raw scores for the 33 students was 27-40. The percentile ranks of the range scores were 29-83rd percentile. Examining the percentile ranks at each level, we find a similarly broad range down to NDRT raw score of 14, which is at the 1st percentile. From these data we concluded that either the NDRT or the DTLS was not accurately measuring the reading comprehension of the students.

In the Fall of 1981 all students who failed to reach the required level on the NDRT were give the *Stanford Diagnostic Reading Test* (SDRT). Table 2 presents the mean and range of scores received by students at each possible NDRT score. The same broad range was found with the SDRT as with the DTLS. We concluded that the NDRT was not accurately measuring the reading comprehension of the students enrolled in our remedial program.

Posttesting at the end of the semester indicated that improvement on the NDRT was not substantiated by SDRT posttest scores. The results on Table 3 indicate that students who scored as low as the 19th percentile on the SDRT were able to score as high as 77th percentile on the NDRT. Conversely, students who scored at the 51st percentile on the SDRT were only able to reach the 17th percentile on the NDRT. We concluded that the posttest score on the NDRT form D was as inaccurate as the pretest score.

Using the 5th stanine of the SDRT as the criteria for adequate reading skills, 40 students who were enrolled in the remedial reading course should not have been. In addition, 33 students who passed the state requirement on the NDRT should not have been passed out of the program.

Cummins (1981, p. 58), like Raygor and Flippo (1981), suggested that appropriate use of the NDRT would be to use it as a "...tool for identifying superior students who could profit from high quality and quantity instructional programs designed for the gifted, as well as useful infor-

5 BEST COPY

mation for reading specialists who need a quick method of identifying students who *may* (his italics) require specific reading remediation after further diagnostic testing." Townsend (1965, p. 1080) made a similar comment about early NDRT forms: "It is not a test which will adequately differentiate among the reading skills of college skills of college students, but it has its place for screening." Raygor (1978) state his feelings more forcefully:

> To use this test on anyone other than highly selected college students or adults will produce very questionable measurement results. Worse than that, it will provide a difficult, defeating, and frustrating experience to many of the students--especially those who have already had enough of such experiences. (pp. 1210-11)

To this author using the NDRT to differentiate between good and poor readers can be compared with using the time it takes persons to drown as a measure of their ability to swim. A re-examination of the use of the NDRT as the instrument which determines whether or not college freshmen possess adequate reading skills is still strongly recommended (Flippo, 1980). That students need special help is a foregone conclusion; that we have targeted the right students is questionable.

Table 1

DTLS Mean and Range Scored by Students at Each NDRT Score.

Pretest Fall 1980

N Students	NDRT RS	NDRT %ile	DTLS RS mean	DTLS mean %ile	DTLS RS range	DTLS range %ile
33	32	22	33.1	47	27-40	28-83
22	30	17	33.6	47	25-40	25-83
27	28	14	31.1	41	21-42	16-83
28	26	11	29.1	36	11-40	1-83
20	24	8	30.5	41	20-40	13-83
8	22	6	29.2	36	24-37	22-66
6	20	4	26.3	36	20-38	13-74
5	18	3	22.4	19	15-35	7-58
4	16	2	25.0	25	18-36	10-66
1	14	1	25.0	25	25	25
2	12	1	20.0	13	15-25	7-25
0	10	1	-	-	-	-
1	8	1	18.0	10	18	10

BEST COPY

Table 2

SDRT Mean and Range Scored by Students at Each NDRT Score.

Pretest Fall 1981

N Students	NDRT RS	NDRT %ile	SDRT RS mean	SDRT mean %ile	SDRT RS range	SDRT range %ile
35	30	17	41.9	35	35-57	9-93
19	28	14	38.4	26	18-54	4-77
17	26	11	38.4	26	16-52	2-68
16	24	8	41.4	33	23-57	7-93
7	22	6	39.1	28	26-55	9-83
6	20	4	38.3	26	31-52	15-68
1	18	3	36.0	23	36	23
3	16	2	31.0	15	23-38	7-28

Table 3

SDRT Mean and Range Scored by Students at Each NDRT Score.

Posttest Fall 1981

N Students	NDRT RS	NDRT %ile	SDRT RS mean	SDRT mean %ile	SDRT RS range	SDRT range %ile
3	52	82	53.3	74	48-55	55-89
8	50	77	46.5	51	35-53	19-79
5	48	72	45.2	44	40-55	29-89
6	46	66	46.3	48	43-51	38-69
10	44	60	43.8	41	28-52	10-74
10	42	54	45.5	44	40-50	29-63
12	40	47	44.7	44	37-52	23-74
7	38	41	43.4	38	32-49	15-59
11	36	35	44.3	41	33-53	16-79
9	34	26	42.1	35	33-54	16-84
6	32	22	41.5	35	33-48	16-55
11	30	17	37.4	23	27-47	9-51
1	28	14	39.0	27	39	27
2	26	11	32.0	15	24-40	7-29
1	24	8	29.0	11	29	11
1	22	6	31.0	13	31	13

References

Brown, J.I., Nelson, M.J. & Denny, E.C. *Examiner's manual: The Nelson-Denny Reading Test* (Forms C and D). Boston: Houghton Mifflin Company, 1976.

Crites, J.O. The Nelson-Denny Reading Test: Vocabulary-Comprehension-Rate. In O.K. Buros (Ed.) *The sixth mental measurements yearbook*. Highland Park, New Jersey: The Gryphon Press, 1965, 1080-1081.

BEST COPY

Cummins, R.P. Test review: The Nelson-Denny Reading Test (Forms E and F). *Journal of Reading*, 1981, 25, 54-59.

Flippo, R.F. The need for comparison studies of college students' reading gains in developmental reading programs using general and specific levels of diagnosis. In M.I. Kamil and A.J. Moe (Eds.), *Perspectives on reading research and instruction*. Twenty-ninth Yearbook of the National Reading Conference. Washington, DC: National Reading Conference, 1980.

Forsyth, R.A. The Nelson-Denny Reading Test (Forms C and D). In O.K. Buros (Ed.), *The eighth mental measurements yearbook* (Vol. 2). Highland Park, New Jersey: The Gryphon Press, 1978, 1207-1208.

Gordon, B., & Flippo, R.F. An update on college reading improvement programs in the southeastern United States. *Journal of Reading*, 1983, 27, 155-163.

Orr, D.B. The Nelson-Denny Reading Test: Vocabulary-Comprehension-Rate. In O.K. Buros (Ed.), *The sixth mental measurements yearbook*. Highland Park, New Jersey: The Gryphon Press, 1965, 1077-1078.

Raygor, A.L. The Nelson-Denny Reading Test (Forms C and D). In O.K. Buros (Ed.) *The eighth mental measurements yearbook* (Vol. 2). Highland Park, New Jersey: The Gryphon Press, 1978, 1209-1211.

Raygor, A.L., & Flippo, R.F. Varieties of comprehension measures: A comparison of intercorrelations among several reading tests. In G.H. McNinch (Ed.), *Comprehension: Process and product*. Hattiesburg, Miss.: American Reading Forum, 1981. (ERIC No. ED 198 485)

Townsend, A. The Nelson-Denny Reading Test: Vocabulary-Comprehension-Rate. In O.K. Buros (Ed.) *The sixth mental measurements yearbook*. Highland Park, New Jersey: The Gryphon Press, 1965, 1078-1080.

Van Blerkom, M.L., & Van Blerkom, D.L. Misleading findings from the Nelson-Denny Reading Test. *Journal of Reading*, 1981, 24, 736-737.